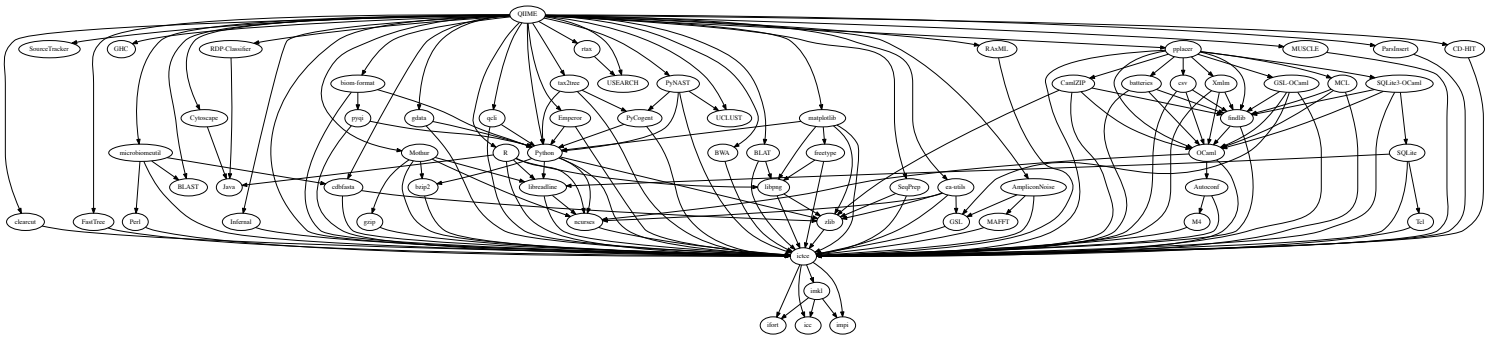
Building (scientific) software with ease

Kenneth Hoste
HPC-UGent, Ghent University, Belgium
kenneth.hoste@ugent.be

lightning talk @ Python devroom - FOSDEM'15
20150131

“Please install QIIME?”

QIIME: Quantitative Insights Into Microbial Ecology (`http://qiime.org/`)



- scientific research domain: bioinformatics . . .
- 59 dependencies in total (*without* compiler toolchain), some optional
  - depends on Haskell (GHC), Java, Python, R, Perl, OCaml, . . .
  - for several deps: **non-standard build procedure** (in various degrees)
- *very* picky about dependency versions (e.g., *must* be Python v2.7.*3*)
- took us several weeks to get it installed (like we wanted). . .
- **. . . now we can (re)build/install it all with a single command!**

(disclaimer: QIIME is not supported yet in the latest EasyBuild release)

EasyBuild in a nutshell (1/2)

EasyBuild is a software build and installation framework that allows you to manage (scientific) software on High Performance Computing (HPC) systems in an *efficient* way, in *collaboration* with other HPC sites.

Requirements:

- Python v2.6 or more recent Python v2.x
- a modules tool (Tcl-based environment modules or Lmod)
- (a system C/C++ compiler)

Terminlogy

The EasyBuild *framework* leverages *easyblocks* to automatically build and install (scientific) software using a particular *compiler toolchain*, as specified by one or multiple *easyconfig files*.

EasyBuild in a nutshell (2/2)

- collection of Python packages and modules, ~50K LOC
- in-house @ HPC-UGent 2009-2012, now GPLv2 (GitHub)
- supports over 550 different software packages

  including CP2K, NAMD, NWChem, OpenFOAM, PETSc, QuantumESPRESSO, WRF, WPS, ...
- *thriving community*: actively contributing, driving development

  Flemish Supercomputer Centre sites, Jülich Supercomputer Centre, Bayer CropScience, UniBas, Cyprus Institute, University of Auckland, Stanford, Texas A&M University, ...
- well documented: `http://easybuild.readthedocs.org`

Building and installing with eb

Build and install WRF (+ 14 dependencies) using Intel compilers & tools:

```
$ eb WRF-3.6.1-intel-2014b-dmpar.eb --robot
== temporary log file in case of crash
/tmp/easybuild-_4lhM4/easybuild-G21y9c.log
== resolving dependencies ...
...
== building and installing HDF5/1.8.13-intel-2014b...
...
== building and installing WRF/3.6.1-intel-2014b...
...
== sanity checking...
== cleaning up...
== creating module...
== COMPLETED: Installation ended successfully
== Results of the build can be found in the log file ...
```

Using the installed software

```
$ module load WRF/3.6.1-intel-2014b
$ which wrf.exe
<prefix>/WRF/3.6.1-intel-2014b-dmpar/WRFV3/main/wrf.exe

$ cat job.sh
#!/bin/bash
module load WRF/3.6.1-intel-2014b
...
mpirun -np 1000 wrf.exe

$ qsub job.sh
```

List of supported software packages

a2ps ABAQUS ABINIT ABySS ACML **ALADIN** Allinea ALLPATHS-LG AMOS AnalyzeFMRI ANSYS ant ANTs APBS ARB argtable aria2 Armadillo arpack-ng ASE ATLAS Autoconf Automake bam2fastq BamTools Bash BayesTraits bbcp bbFTP bbftpPRO bc beagle-lib Beast BEDOPS BEDTools BFAST binutils BioPerl Biopython BiSearch Bison BitSeq BLACS BLAST BLAT BOINC Bonnie++ Boost Bowtie Bowtie2 BWA byacc bzip2 cairo CAP3 CBLAS ccache CCfits CD-HIT CDO CEM CFITSIO cflow CGAL cgdb Chapel CHARMM Chimera Circos Clang CLHEP CLooG Clustal-Omega ClustalW2 CMake Coreutils Corkscrew **CP2K** CPLEX CRF++ ctffind Cube CUDA Cufflinks cURL cutadapt CVS CVXOPT Cython DB DBD-mysql DBD-SQLite DB_File DIALIGN-TX Diffutils DL_POLY_Classic Docutils **DOLFIN** Doxygen **EasyBuild** ECore ed Eigen ELinks ELPA ELPH Emacs EMBOSS EPD ErlangOTP ESMF ESPResSo evmix expat eXpress FASTA fastahack FastTree FASTX-Toolkit FCM FDS FDTD_Solutions Ferret FFC ffmpeg FFTW FIAT file findutils fixesproto flex FLTK FLUENT fmri FoldX fontconfig FRC_align freeglut FreeSurfer freetype FSL g2clib g2lib GAMESS-US GATE GATK gawk GCC GD GDAL GDB Geant4 GEM-library GEMSTAT GenomeAnalysisTK GEOS gettext GHC Ghostscript GIMPS git GLib GLIMMER GLPK glproto GMAP-GSNAP GMP GMT gnuplot gnutls Go GObject-Introspection google-sparsehash GPAW gperf gperftools grace Graphviz GraphViz2 Greenlet grep grib_api GROMACS GSL gsl GSSAPI GTI GTS guile gzip h4toh5 h5py h5utils HarfBuzz Harminv HDF HDF5 HH-suite HMMER horton HPCG HPL HTSeq HTSlib hwloc Hypre icc ifort imake imkl impi Infernal inputproto Inspector Instant intltool iompi IOR Iperf ipp IPython IsoInfer ispc itac JAGS Jansson JasPer Java Jellyfish Jinja2 JUnit kbproto Kerberos_V5 LAPACK less lftp libcap-ng libcircle libctl libdrm libevent libffi libgd libgtextutils libharu libibmad libibumad libibverbs libICE libidn Libint libint2 libjpeg-turbo libmatheval libpciaccess libpng libpthread-stubs libreadline libSM libsmm LIBSVM LibTIFF libtool libudev libungif libunistring libunwind libX11 libXau libXaw libxc libxcb libXdmcp libXext libXfixes libXft libXi libXinerama libxml2 libXmu libXp libXpm libXrender libxslt libXt libyaml likwid Lmod Lua LWM2 lxml lynx LZO M4 MAFFT make makedepend Maple MariaDB Mathematica MATLAB matplotlib Maven mc MCL mcpp MDP mdtest Meep MEME Mercurial Mesa Mesquite MetaVelvet MethPipe METIS MMSEQ Modeller Molden Molekel molmod Mothur motif MPFR mpi4py mpiBLAST MPICH MPICH2 MrBayes MTL4 MUMmer MUMPS MUSCLE MUST MUSTANG MVAPICH2 MySQL NAMD nano NASM NCBI-Toolkit ncdf4 **NCL** ncurses ncview NEdit netaddr netCDF netCDF-C++ netCDF-C++4 netCDF-Fortran netcdf4-python netifaces NetLibIDN netloc nettle NEURON nodejs ns numactl numexpr numpy NWChem O2scl Oases OCaml Oger OPARI2 OpenBabel OpenBLAS OpenCV **OpenFOAM OpenFOAM-Extend** OpenIFS OpenMD OpenMPI OpenPGM OpenSees OpenSSL ORCA orthomcl otcl OTF OTF2 packmol PAML pandas PANDAseq Pango PAPI parallel Paraview ParFlow ParMETIS ParMGridGen Pasha patch paycheck PCC PCRE PDT Perl **PETSc** petsc4py phonopy PhyML picard pixman pkg-config PLINK PnMPI popt PP PRANK Primer3 printproto problog protobuf pscom PSI psmpi psmpi2 PyQuante pysqlite pyTables **Python** python-dateutil python-meep PyYAML PyZMQ Qhull QLogicMPI Qt qtop QuadProg++ **QuantumESPRESSO** R RAxML RCS RDP-Classifier RELION renderproto rjags RNAz ROOT Rosetta rSeq RSEQtools Ruby runjags Sablotron SAMtools ScaLAPACK Scalasca ScientificPython scikit-learn scipy SCons SCOOP Score-P SCOTCH SDCC SDPA sed segemehl setuptools Shapely SHRiMP SIBELia sickle Silo slalib-c SLEPc SOAPaligner SOAPdenovo SOAPdenovo2 SOAPec SPAdes Sphinx SPRNG SQLite SRA-Tookit Stacks stemming Stow Stride SuiteSparse SURF SWIG sympy systemd Szip TAMkin Tar tbb TCC Tcl tclcl tcsh Tesla-Deployment-Kit texinfo Theano TiCCutils TiMBL TinySVM Tk TopHat Tornado TotalView TREE-PUZZLE Trilinos Trinity UDUNITS UFC UFL util-linux Valgrind VCFtools Velvet ViennaRNA Vim Viper vsc-base vsc-mympirun vsc-mympirun-scoop vsc-processcontrol VSC-tools VTK VTune WHAM **WIEN2k** wiki2beamer **WPS WRF** xbitmaps xcb-proto XCrySDen xextproto xineramaproto XML XML-Dumper XML-LibXML XML-Parser XML-Simple XML-Twig xorg-macros xproto xtrans XZ yaff YamCha YAML-Syck Yasm YAXT ZeroMQ zlib zsh zsync

Do you want to know more?

- website: `http://hpcugent.github.com/easybuild`
- documentation: `http://easybuild.readthedocs.org`
- stable releases @ PyPi `pypi.python.org/pypi/easybuild`
- sources @ GitHub: `https://github.com/hpcugent/easybuild`*
- EasyBuild mailing list: easybuild@lists.ugent.be
  `https://lists.ugent.be/wws/subscribe/easybuild`
- IRC: `#easybuild` on `chat.freenode.net`
- Twitter: @easy_build